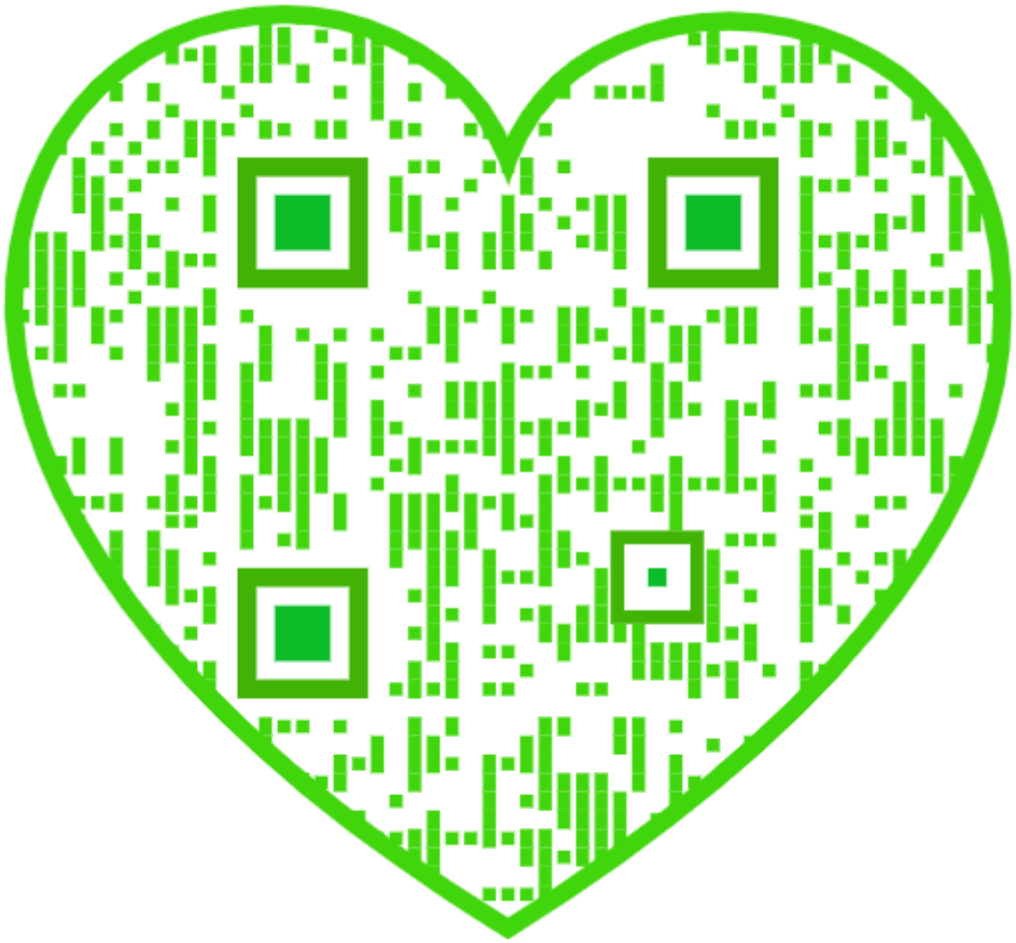
Master in Artificial Intelligence

Ethical Considerations II



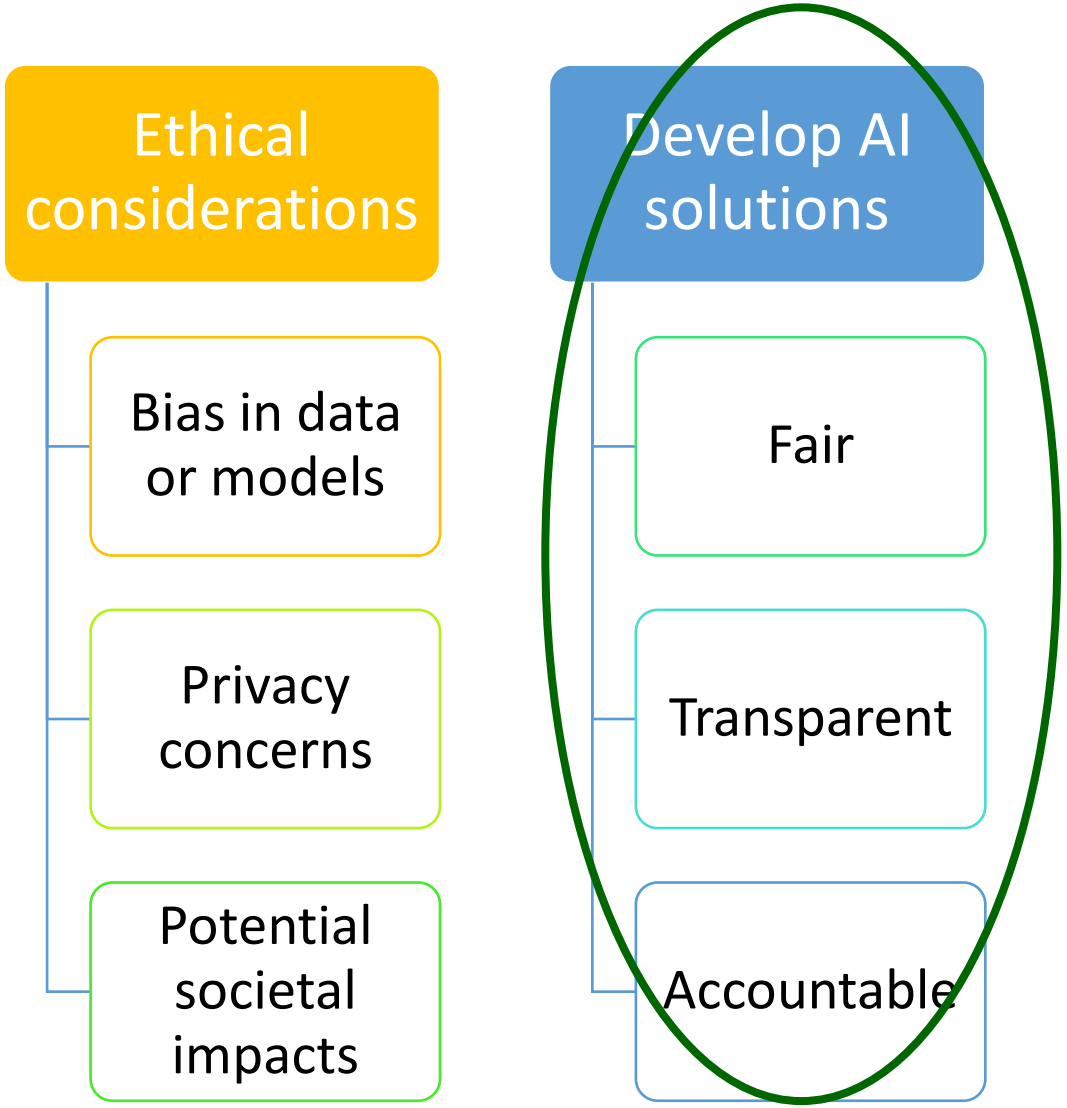
Purpose

The purpose of the section is to help you learn how to take care of ethical considerations in artificial intelligence and machine learning to become a Successful Artificial Intelligence (AI) Engineer

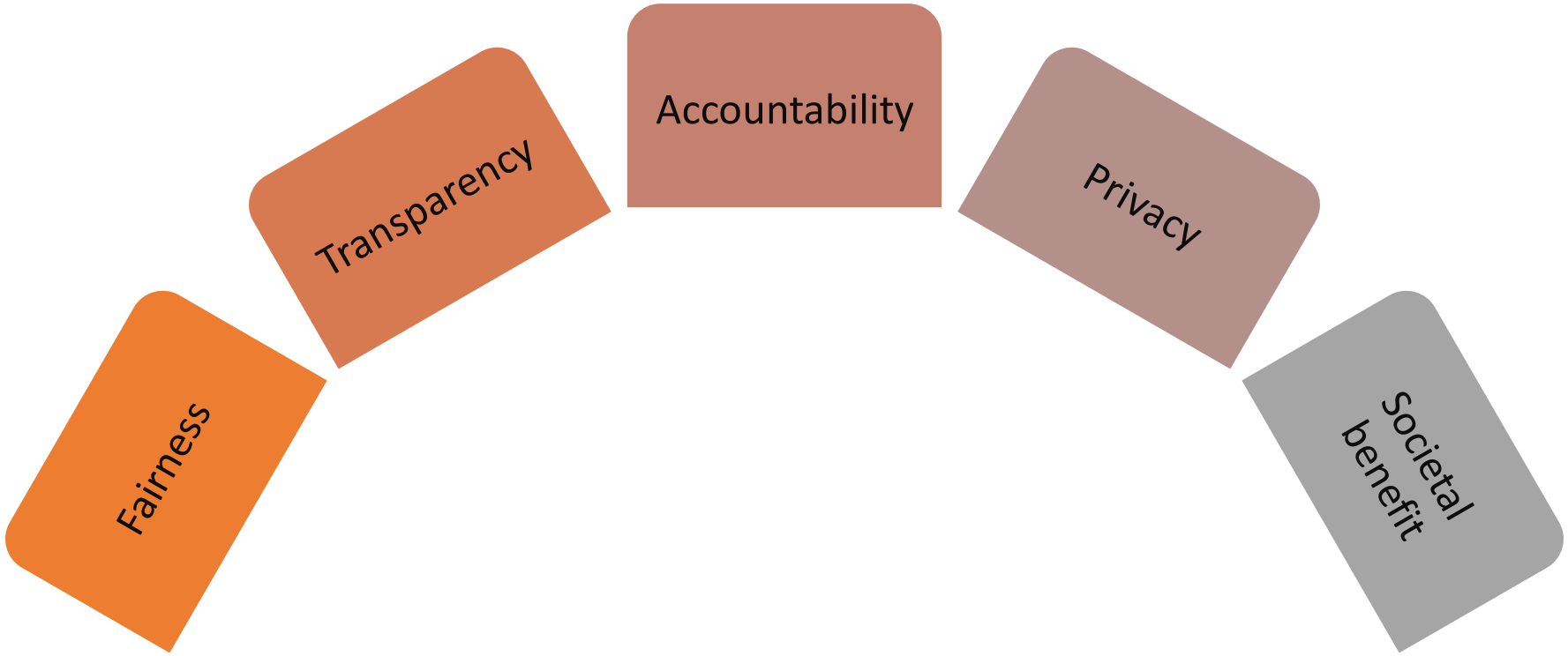
At the end of this lecture, you will learn the following

- **How to develop AI solutions that are fair, transparent, and accountable**

How to develop AI solutions that are fair, transparent, and accountable



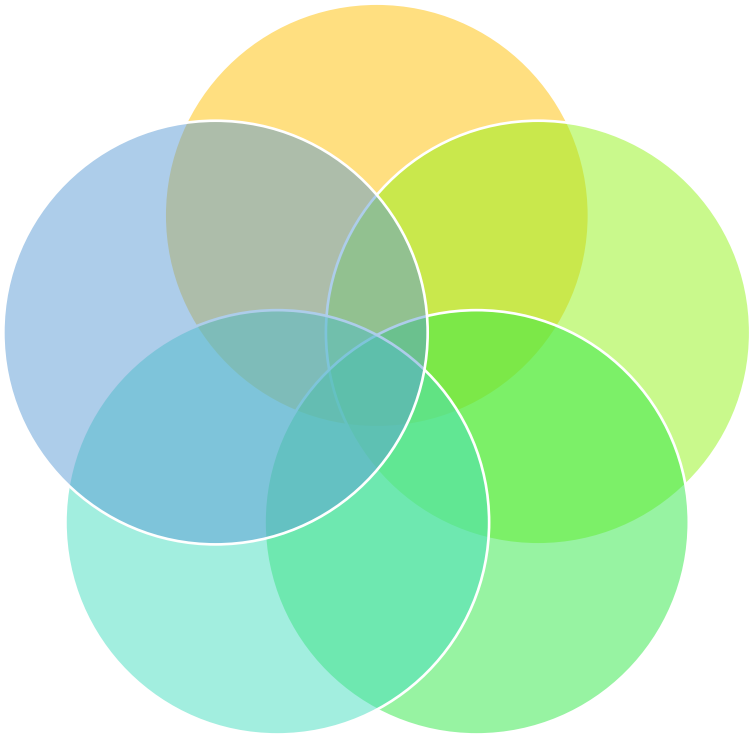
Define Ethical Principles



Diverse and Inclusive Teams

Data scientists

Stakeholders from affected communities



Software engineers

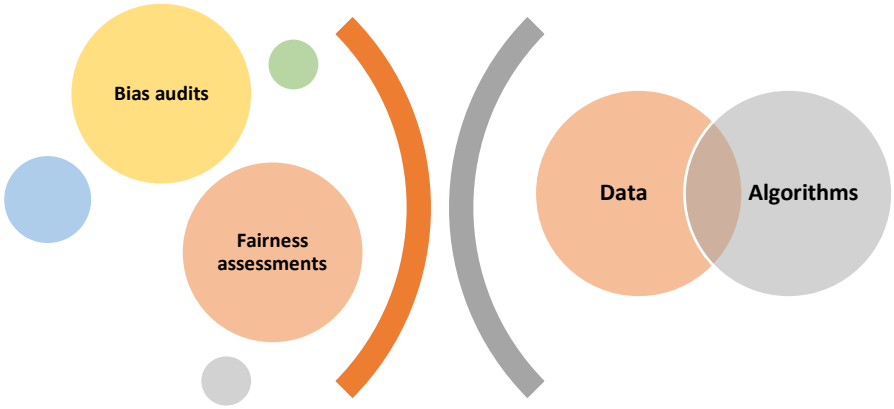
Ethicists

Domain experts

Ethical Design and Development

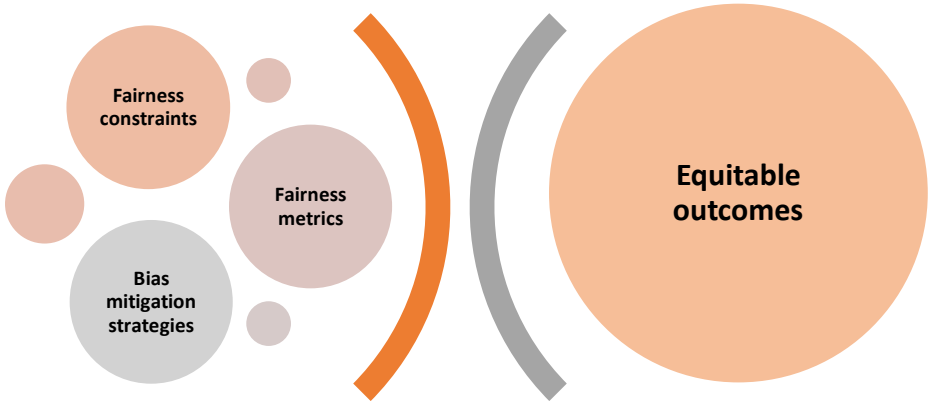
- **Ethical Design Canvas**
  - Stakeholders
  - Impact Zones
  - Ethical Risks
  - Ethical Principles
  - Mitigation Strategies
- **Value Sensitive Design (VSD)**
  - Identify Values
  - Analyze Stakeholders
  - Incorporate Values into Design
  - Evaluate Trade-offs
  - Iterative Refinement
- **Protect user privacy and data right**
  - Privacy-preserving techniques
  - Data anonymization
  - Differential privacy to

Fairness and Bias Mitigation



Conduct

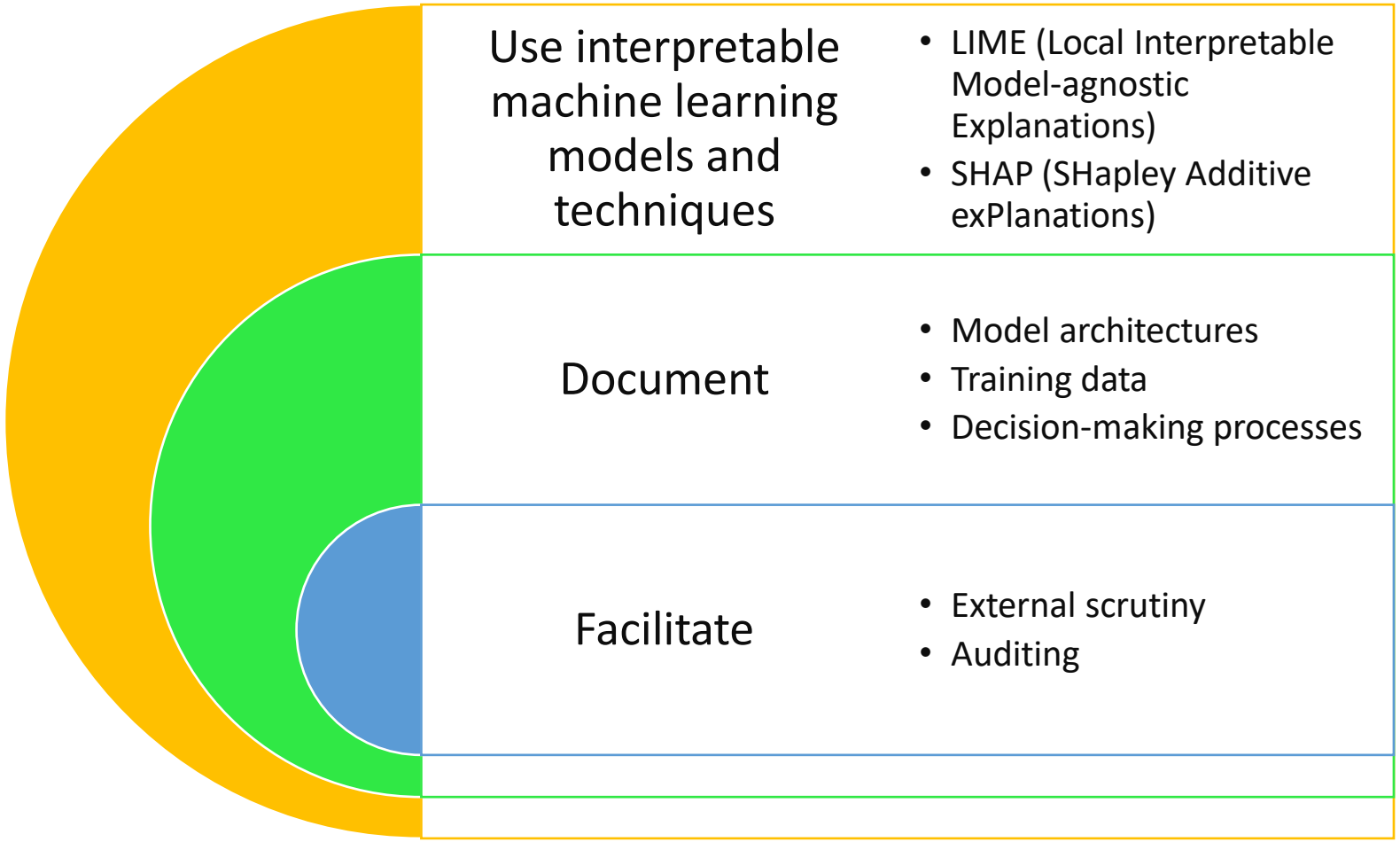
Identify and mitigate biases



Implement fairness-aware machine learning techniques

Promote

Transparency and Explainability



Accountability and Oversight

Establish mechanisms

- For accountability and oversight
- Throughout the AI development lifecycle.

Implement governance structures

- Such as ethics committees or review boards
- To review and assess the ethical implications of AI projects.

Define roles and responsibilities

- Decision-making
- Monitoring
- Addressing ethical concerns

User Empowerment and Consent

Empower users

- Providing them with control over their data
- Ability to understand and manage their interactions with AI systems.

Obtain informed consent from users

- Data collection
- Processing
- Algorithmic decision-making.

Provide mechanisms for users

- Access
- Review
- Correct
- Data and algorithmic decisions

Continuous Evaluation and Improvement

Continuously evaluate AI solutions for

- Fairness
- Transparency
- Accountability post-deployment

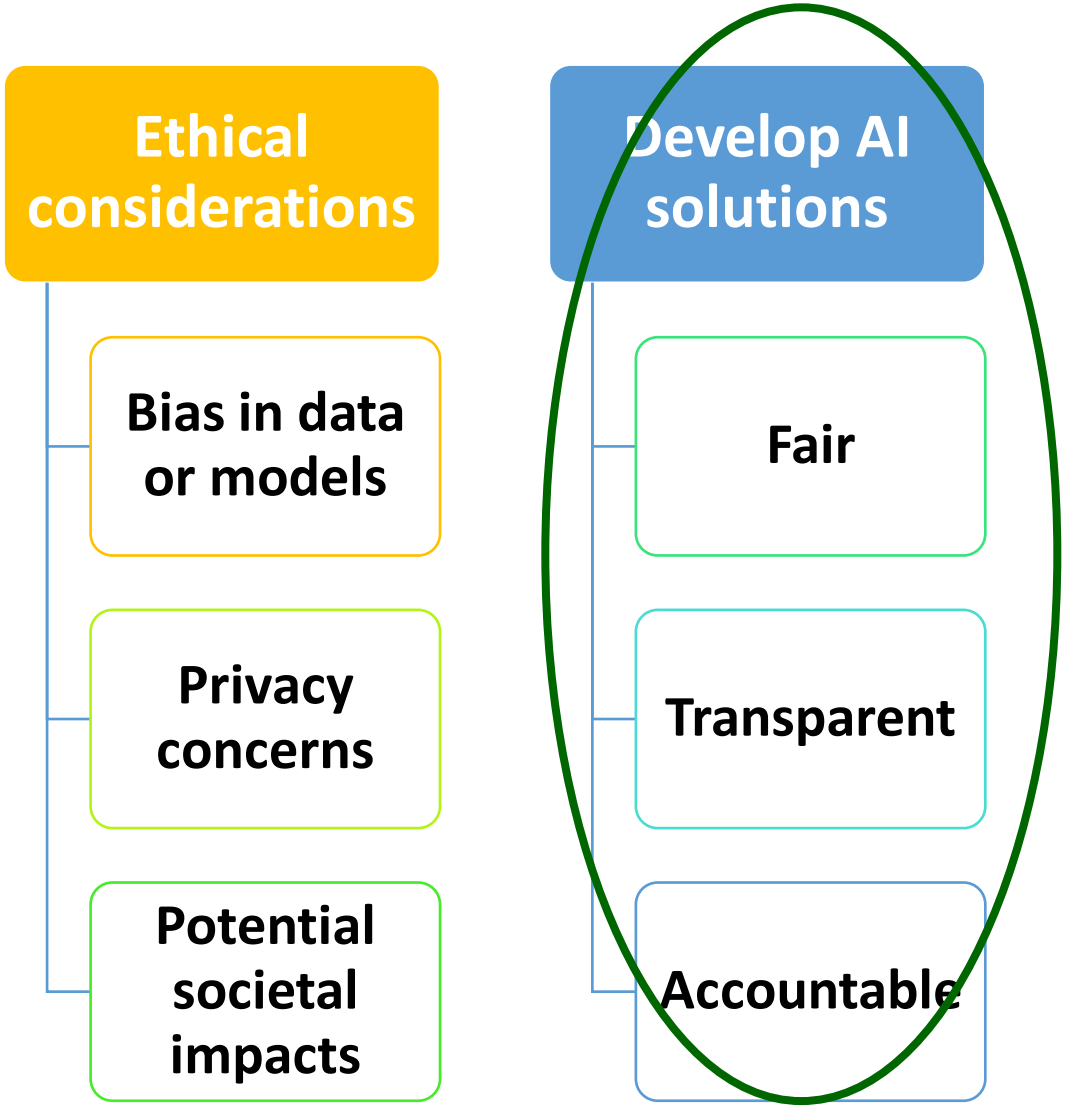
Solicit feedback from

- Stakeholders and affected communities
- Identify areas for improvement
- Address emerging ethical issues

Iterate on AI solutions based on

- Feedback
- New insights
- Evolving ethical standards

How to develop AI solutions that are fair, transparent, and accountable



What is next?

An example of taking care of ethical considerations in artificial intelligence and machine learning

Ethical considerations	Develop AI solutions
Bias in data or models	Fair
Privacy concerns	Transparent
Potential societal impacts	Accountable

Master in Artificial Intelligence

Ethical Considerations II